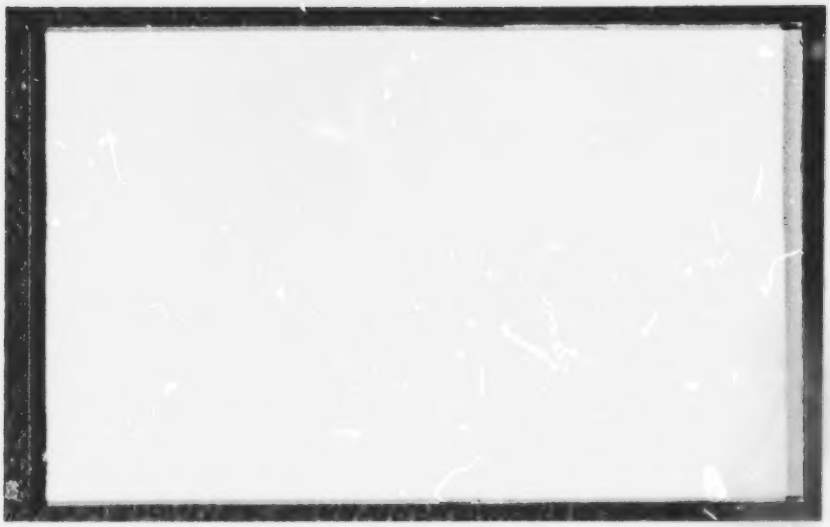
UNIVERSITÉ DE MONTRÉAL



CENTRE DE RECHERCHE SUR LES TRANSPORTS

CASE POSTALE 6128, SUCCURSALE A, MONTRÉAL, QUÉBEC, H3C 3J7

HETEROSCEDASTICITY AND THE USE OF BOX-COX TRANSFORMATIONS

by

Marc J.I. Gaudry and Marcel G. Dagenais

The Research and Development Centre of Transport Canada and the F.C.A.C. program of the *Ministère de l'Éducation du Québec* have supported the research on which this paper is based. The authors thank T.C. Liem for comments on an early version.

This paper appeared in Economics Letters, Vol. 2, No. 3, pp. 225-229, 1979.

Université de Montréal
Centre de recherche sur les transports - Publication # 106*
1979

ABSTRACT

Direct and inverse power Box-Cox transformations are used to characterize the form of heteroscedasticity in models where the variables themselves are subjected to direct Box-Cox transformations. Resulting formulations are extended to take serial correlation simultaneously into account.

1. Introduction

Monotonic transformations studied by Anscombe and Tukey (1954), Box and Tidwell (1962) and Box and Cox (1964) have been increasingly used in regression models. Typically the form of the "fixed part" of the model is analyzed independently from that of the "stochastic part" of the model. Recently, some authors have begun to explore the structure of the stochastic part simultaneously with the functional forms of the fixed part: Savin and White (1978) have considered functional form and first order autocorrelation of residuals and Gaudry and Wills (1978) have extended the approach to higher order autocorrelation. Despite the presumption that Box-Cox transformation parameters may compensate for heteroscedasticity and be very sensitive to its eventual presence [e.g. Zarembka (1974)], we are not aware of general procedures which would explicitly take this problem into account. In this letter, we shall approach both the issue of the analytical form of hererо-scedasticity and the issue of its estimation simultaneously with that of the functional form of the fixed part of the model; we will then briefly extend the analysis to incorporate autocorrelation.

The model written in the usual notation for each observation is

$$y_t^{(\lambda_y)} = \sum_{k=1}^{K} \beta_k X_{k_t}^{(\lambda_{xk})} + u_t \quad , \tag{1}$$

$$u_t = [f(Z_{1_t}, \ldots, Z_{M_t})]^{1/2} v_t \quad , \qquad t=1, \ldots, N, \tag{2}$$

where the explanatory variables X and Z may belong to different subsets; furthermore, $E(u)=0$ and $E(uu')=\Omega$ with representative element

$$\omega_{tt'} = \begin{cases} \Psi^2 f(Z_{1_t}, \ldots, Z_{M_t}) , & t=t', \\ 0, & t \neq t', \end{cases} \tag{3}$$

where $u_.$ is the N x 1 vector of u_t elements ($t=1, \ldots, N$) and Ω is an N x N matrix.

It is also the case that $E(v_.) = 0$ and, if I denotes the identity matrix,

$$E(v\, v') = \Psi^2 I , \tag{4}$$

where again v is N x 1. Moreover, the β_k are parameters and the Box-Cox transformation of any strictly positive non-Boolean variable w is defined for the λ_w parameter as

$$w^{(\lambda_w)} \begin{cases} = \dfrac{w^{\lambda_w} - 1}{\lambda_w} , & \lambda_w \neq 0 , \\ = \ell n\, w , & \lambda_w = 0 , \end{cases} \tag{5}$$

with the corresponding inverse function:

$$w = \begin{cases} (\lambda_w\, w^{(\lambda_w)} + 1)^{1/\lambda_w} , & \lambda_w \neq 0, \\ \exp(w^{(\lambda_w)}) , & \lambda_w = 0 . \end{cases} \tag{6}$$

2. The form of heteroscedasticity

Assume that, in (3), any given diagonal element of Ω may be expressed as

$$\omega_{tt} = \Psi^2 \{\lambda_u [\delta_0 + \sum_{m=1}^{M} \delta_m Z_{m_t}^{(\lambda_{em})}] + 1\}^{1/\lambda_u} \tag{7}$$

in which direct Box-Cox transformations are applied to the Z_m, the variables which explain error variance, and an inverse power transformation is applied to the expression in the square parentheses. The constants δ_0 and Ψ are necessary to preserve (Schlesselman, 1971) the invariance of the transformation to units of measurement of the Z_m and y. Equation (7) contains a great number of special cases of traditional interest.

Setting $\lambda_u = 0$, yields, by (6),

$$\tilde{\omega}_{tt} = \Psi^2 \exp(\delta_0 + \sum_m \delta_m Z_{m_t}^{(\lambda_{em})}) \equiv \sigma^2 [\exp (\sum_m \delta_m Z_{m_t}^{(\lambda_{em})})] , \qquad (8)$$

where $\sigma^2 = \Psi^2 \exp(\delta_0)$. This reduces to the multivariate multiplicative model

$$\tilde{\tilde{\omega}}_{tt} = \sigma^2 \prod_m Z_{m_t}^{\delta_m} \qquad (9)$$

found in Dagum and Dagum (1974) when each λ_{em} is set equal to 0; it reduces further to the classical multivariate form

$$\tilde{\tilde{\tilde{\omega}}}_{tt} = \sigma^2 \prod_m Z_{m_t}^2 \qquad (10)$$

when each δ_m is set equal to 2. Park's (1966) specification is obtained by setting all δ_m but one equal to zero in (8) and by setting the λ_{em} of the remaining variable equal to 0. Imposing the further constraint that the retained $\delta_m = 2$ yields the classical univariate form $\sigma^2 Z_m^2$.

Another family of special cases is obtained by setting $\lambda_u = 1$; then,

$$\bar{\bar{\omega}}_{tt} = \bar{\bar{\delta}}_0 + \sum_m \bar{\bar{\delta}}_m Z_{m_t}^{(\lambda_{em})} , \qquad \bar{\bar{\delta}}_0 = \Psi^2 (\delta_0 + 1) , \quad \bar{\bar{\delta}}_m = \Psi^2 \delta_m , \qquad (11)$$

which includes the linear form (all $\lambda_{em} = 1$) and many nonlinear forms (e.g.

square root if the $\lambda_{em} = 1/2$ and square power if the $\lambda_{em} = 2$) a special cases.

The subclass defined by (8) is more useful then the subclass defined by (11) because the latter can produce negative elements in Ω. Form (8) also leaves considerable room for interaction among the variables most likely to explain error variance. We shall use form (8) for expository purposes in the following sections.

3. The likelihood function with heteroscedasticity

If, in (1), the error term is assumed to be distributed normally with zero mean and covariance matrix $\tilde{\Omega}$ explicated in (8), the likelihood function is

$$L=\left(\frac{1}{\sqrt{2\pi}}\right)^{N}|\tilde{\Omega}|^{-1/2}\exp\left\{-\frac{1}{2}(y^{(\lambda_y)}-X^{(\lambda_x)}\beta)'\,\tilde{\Omega}^{-1}\,(y^{(\lambda_y)}-X^{(\lambda_x)}\beta)\right\}|J(\lambda_y;y)|\,, \qquad (12)$$

where model (1) is written in matrix notation and $|J(\lambda_y;y)|$ is the Jacobian of the transformation from $y_t^{(\lambda_y)}$ to the actually observed y_t:

$$|J(\lambda_y;y)| = \left|\det\left[\frac{\partial y_t^{(\lambda_y)}}{\partial y_t}\right]\right| = \prod_{t=1}^{n} y_t^{\lambda_y-1}\,. \qquad (13)$$

Equation (12) can be rewritten as

$$L = \prod_{t=1}^{N} \frac{y_t^{\lambda_y-1}}{\sqrt{2\pi}\,\sigma\sqrt{\exp(\sum_m \delta_m Z_{m_t}^{(\lambda_{em})})}}\exp\left[-\frac{(y_t^{(\lambda_y)}-\sum_k \beta_k X_{k_t}^{(\lambda_{xk})})^2}{2\sigma^2\exp(\sum_m \delta_m Z_{m_t}^{(\lambda_{em})})}\right], \qquad (14)$$

from which one may estimate parameters β_k, σ, δ_m, λ_y, λ_{xk} and λ_{em}.

In terms of testing for the form of heteroscedasticity jointly with the form of the fixed part of the model, one approach is to obtain asymptotic variances and covariances for these parameters from the matrix of second partial derivatives of the likelihood function and to perform asymptotic tests based on the normal distribution. Another approach is to use the likelihood ratio test and to compare the value of (14) at the maximum under more or less severe constraints. It is therefore possible to deal analytically with a large number of structures without confounding problems of functional form and of heteroscedasticity.

4. Extension to consider autocorrelation simultaneously

To model specification (1) - (2), we may add a stationary first order autocorrelation process

$$v_t = \rho_1 v_{t-1} + s_t \,, \tag{15}$$

where s_t is normally and independently distributed with mean zero and variance $\psi^{*2} = \psi^2 (1 - \rho_1^2)$. In the special case of equation (8), the corresponding expression would be $\sigma^{*2} = \sigma^2(1 - \rho_1^2)$. Expression (14) then becomes

$$L = (1-\rho_1^2)^{1/2} \frac{y_1^{\lambda_y - 1}}{\sqrt{2\pi}\,\sigma^* \sqrt{\exp\left(\sum_m \delta_m Z_{m_1}^{(\lambda_{em})}\right)}} \exp\left[- \frac{(1-\rho_1)^2 \left(y_1^{(\lambda_y)} - \sum_k \beta_k X_{k_1}^{(\lambda_{xk})}\right)^2}{2\sigma^{*2} \exp\left(\sum_m \delta_m Z_{m_1}^{(\lambda_{em})}\right)} \right] \cdot$$

$$\prod_{t=2}^{N} \frac{y_t^{\lambda_y - 1}}{\sqrt{2\pi}\,\sigma^* \sqrt{\exp\left(\sum_m \delta_m Z_{m_t}^{(\lambda_{em})}\right)}} \exp\left[- \frac{1}{2\sigma^{*2}} \left[\frac{\left(y_t^{(\lambda_y)} - \sum_k \beta_k X_{k_t}^{(\lambda_{xk})}\right)}{\sqrt{\exp\left(\sum_m \delta_m Z_{m_t}^{(\lambda_{em})}\right)}} - \rho_1 \frac{\left(y_{t-1}^{(\lambda_y)} - \sum_k \beta_k X_{k_{t-1}}^{(\lambda_{xk})}\right)}{\sqrt{\exp\left(\sum_m \delta_m Z_{m_{t-1}}^{(\lambda_{em})}\right)}} \right]^2 \right] , \tag{16}$$

where the component which refers to the first observation has been isolated.

Further details and a simplified formulation which drops the first observations in the presence of higher order autocorrelation are given in a working paper (Gaudry and Dagenais, 1978). A program is available to compute the maximum likelihood solution of equation (16) in the presence of higher order autocorrelation, with as many as five non zero autocorrelation parameters (Liem, 1979); the program also evaluates the asymptotic covariance matrix of the parameter estimates.

REFERENCES

. Anscombe, F.J. and Tukey, J.W., 1954, The criticism of transformation. Unpublished paper presented before the American Statistical Association and Biometric Societies, Montreal.

. Box, G.E.P. and Cox, D.R., 1964, An analysis of transformations. Journal of the Royal Statistical Society B, 211-243.

. Box, G.E.P. and Tidwell, P.W., 1962, Transformation of the independent variables, Technometrica 4, 531-550.

. Dagum, C. et Dagum, E.B., 1974, Construction de modèles et analyse économétrique. Cahiers de l'Institut des sciences mathématiques et économiques appliquées 8, No. 11-12.

. Gaudry, M.J.I. and Dagenais, M.G., 1978, The use of Box-Cox transformations in regression models with heteroscedastic autoregressive residuals, Cahiers 7814, Département de sciences économiques, Université de Montréal.

. Gaudry, M.J.I. and Wills, M.J., 1978, Estimating the functional form of travel demand models. Transportation Research 12, 257-289.

. Liem, T.C., 1979, L-1: a program for Box-Cox transformations in regression models with heteroskedastic and autoregressive residuals. Publication # 134 , Centre de recherche sur les transports, Université de Montréal.

. Park, R.E., 1966, Estimation with heteroskedastic error terms. Econometrica 34, 888.

. Savin, N.E. and White, K.J., 1978, Estimation and testing for functional form and autocorrelation: a simultaneous approach. Journal of Econometrics 8, 1-12.

. Schlesselmann, J., 1971, Power families: a note on the Box and Cox transformation. Journal of the Royal Statistical Society B, 33, 307-311.

. Zarembka, P., 1974, Transformation of variables in econometrics, in P. Zarembka, ed., Frontiers in Econometrics, ch. 3.